

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/820,745

DATE: 11/26/2001

TIME: 12:34:34

Input Set : A:\620-139.app

Output Set: N:\CRF3\11212001\I820745.raw

ENTERED

```

3 <110> APPLICANT: Blundell, Tom L
4      Abell, Christopher
5      Inoue, Tsuyoshi
6      von Delft, Frank
8 <120> TITLE OF INVENTION: Crystal Structure
10 <130> FILE REFERENCE: 620-139
12 <140> CURRENT APPLICATION NUMBER: US 09/820,745
13 <141> CURRENT FILING DATE: 2001-03-30
15 <160> NUMBER OF SEQ ID NOS: 12
17 <170> SOFTWARE: PatentIn Ver. 2.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 8
21 <212> TYPE: PRT
22 <213> ORGANISM: Artificial Sequence
24 <220> FEATURE:
25 <223> OTHER INFORMATION: Description of Artificial Sequence: Conserved
26      sequence motif
28 <400> SEQUENCE: 1
29 Leu Val Gly Asp Ser Leu Gly Met
30 1      5
33 <210> SEQ ID NO: 2
34 <211> LENGTH: 6
35 <212> TYPE: PRT
36 <213> ORGANISM: Artificial Sequence
38 <220> FEATURE:
39 <223> OTHER INFORMATION: Description of Artificial Sequence: Conserved
40      sequence motif
42 <400> SEQUENCE: 2
43 Val Lys Ile Glu Gly Gly
44 1      5
47 <210> SEQ ID NO: 3
48 <211> LENGTH: 8
49 <212> TYPE: PRT
50 <213> ORGANISM: Artificial Sequence
52 <220> FEATURE:
53 <223> OTHER INFORMATION: Description of Artificial Sequence: Conserved
54      sequence motif
56 <220> FEATURE:
57 <221> NAME/KEY: SITE
58 <222> LOCATION: (3)
59 <223> OTHER INFORMATION: Xaa is a hydrophobic residue
61 <400> SEQUENCE: 3
W--> 62 Gly His Xaa Gly Leu Thr Pro Gln
63 1      5
66 <210> SEQ ID NO: 4
67 <211> LENGTH: 7
68 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/820,745

DATE: 11/26/2001

TIME: 12:34:34

Input Set : A:\620-139.app

Output Set: N:\CRF3\11212001\I820745.raw

```

69 <213> ORGANISM: Artificial Sequence
71 <220> FEATURE:
72 <223> OTHER INFORMATION: Description of Artificial Sequence: Conserved
73     sequence motif
75 <400> SEQUENCE: 4
76 Gly Gly Tyr Lys Val Gln Gly
77   1           5
80 <210> SEQ ID NO: 5
81 <211> LENGTH: 6
82 <212> TYPE: PRT
83 <213> ORGANISM: Artificial Sequence
85 <220> FEATURE:
86 <223> OTHER INFORMATION: Description of Artificial Sequence: Conserved
87     sequence motif
89 <400> SEQUENCE: 5
90 Ile Gly Ile Gly Ala Gly
91   1           5
94 <210> SEQ ID NO: 6
95 <211> LENGTH: 6
96 <212> TYPE: PRT
97 <213> ORGANISM: Artificial Sequence
99 <220> FEATURE:
100 <223> OTHER INFORMATION: Description of Artificial Sequence: Conserved
101     sequence motif
103 <400> SEQUENCE: 6
104 Asp Gly Asn Ile Leu Val
105   1           5
108 <210> SEQ ID NO: 7
109 <211> LENGTH: 264
110 <212> TYPE: PRT
111 <213> ORGANISM: Escherichia coli
113 <400> SEQUENCE: 7
114 Met Lys Pro Thr Thr Ile Ser Leu Leu Gln Lys Tyr Lys Gln Asp Lys
115   1           5           10           15
117 Lys Arg Phe Ala Thr Ile Thr Ala Tyr Asp Tyr Ser Phe Ala Lys Leu
118           20           25           30
120 Phe Ala Asp Glu Gly Leu Asn Val Met Leu Val Gly Asp Ser Leu Gly
121           35           40           45
123 Met Thr Val Gln Gly His Asp Ser Thr Leu Pro Val Thr Val Ala Asp
124           50           55           60
126 Ile Ala Tyr His Thr Ala Ala Val Arg Arg Gly Ala Pro Asn Cys Leu
127   65           70           75           80
129 Leu Leu Ala Asp Leu Pro Phe Met Ala Tyr Ala Thr Pro Glu Gln Ala
130           85           90           95
132 Phe Glu Asn Ala Ala Thr Val Met Arg Ala Gly Ala Asn Met Val Lys
133           100          105          110
135 Ile Glu Gly Gly Glu Trp Leu Val Glu Thr Val Gln Met Leu Thr Glu
136           115          120          125
138 Arg Ala Val Pro Val Cys Gly His Leu Gly Leu Thr Pro Gln Ser Val

```

RAW SEQUENCE LISTING

DATE: 11/26/2001

PATENT APPLICATION: US/09/820,745

TIME: 12:34:34

Input Set : A:\620-139.app

Output Set: N:\CRF3\11212001\I820745.raw

```

139      130      135      140
141 Asn Ile Phe Gly Gly Tyr Lys Val Gln Gly Arg Gly Asp Glu Ala Gly
142 145      150      155      160
144 Asp Gln Leu Leu Ser Asp Ala Leu Ala Leu Glu Ala Ala Gly Ala Gln
145      165      170      175
147 Leu Leu Val Leu Glu Cys Val Pro Val Glu Leu Ala Lys Arg Ile Thr
148      180      185      190
150 Glu Ala Leu Ala Ile Pro Val Ile Gly Ile Gly Ala Gly Asn Val Thr
151      195      200      205
153 Asp Gly Gln Ile Leu Val Met His Asp Ala Phe Gly Ile Thr Gly Gly
154      210      215      220
156 His Ile Pro Lys Phe Ala Lys Asn Phe Leu Ala Glu Thr Gly Asp Ile
157 225      230      235      240
159 Arg Ala Ala Val Arg Gln Tyr Met Ala Glu Val Glu Ser Gly Val Tyr
160      245      250      255
162 Pro Gly Glu Glu His Ser Phe His
163      260
166 <210> SEQ ID NO: 8
167 <211> LENGTH: 267
168 <212> TYPE: PRT
169 <213> ORGANISM: Schizosaccharomyces pombe
171 <400> SEQUENCE: 8
172 Met Ser Leu Lys Gln Ile Thr Ile Ser Thr Leu Arg Gln Trp Lys Leu
173 1 5 10 15
175 Ala Asn Lys Lys Phe Ala Cys Ile Thr Ala Tyr Asp Ala Ser Phe Ser
176 20 25 30
178 Arg Leu Phe Ala Glu Gln Gly Met Pro Val Met Leu Val Gly Asp Ser
179 35 40 45
181 Leu Gly Met Thr Ala Gln Gly His Ser Thr Thr Leu Pro Val Ser Val
182 50 55 60
184 Glu Asp Ile Ala Tyr His Thr Lys Ser Val Arg Arg Gly Ala Pro Asn
185 65 70 75 80
187 Arg Leu Leu Met Ala Asp Leu Pro Phe Met Ser Tyr Ser Thr Trp Glu
188 85 90 95
190 Asp Ala Cys Lys Asn Ala Ala Thr Val Met Arg Ala Gly Ala Asn Ile
191 100 105 110
193 Val Lys Ile Glu Gly Gly Gly Asn Trp Ile Phe Glu Ile Val Gln Arg
194 115 120 125
196 Leu Thr Glu Arg Ser Val Pro Val Ala Gly His Leu Gly Leu Thr Pro
197 130 135 140
199 Gln Ser Val Asn Ile Phe Gly Gly Tyr Lys Ile Gln Gly Arg Glu Gln
200 145 150 155 160
202 Ser Ala Ala Ala Arg Leu Ile Glu Asn Ala Gln Gln Leu Glu Lys Phe
203 165 170 175
205 Gly Ala Gln Leu Leu Val Leu Glu Cys Ile Pro Glu Ser Leu Ala Glu
206 180 185 190
208 Gln Ile Thr Lys Thr Ile Ser Ile Pro Thr Ile Gly Ile Gly Ala Gly
209 195 200 205
211 Lys His Thr Asp Gly Gln Ile Leu Val Met His Asp Ala Leu Gly Ile

```

RAW SEQUENCE LISTING

DATE: 11/26/2001

PATENT APPLICATION: US/09/820,745

TIME: 12:34:34

Input Set : A:\620-139.app

Output Set: N:\CRF3\11212001\I820745.raw

```

212      210      215      220
214 Thr Gly Gly Arg Pro Pro Lys Phe Ala Lys Asn Phe Leu Ser Gly Ala
215 225      230      235      240
217 Gly Asp Ile Arg Thr Ala Ile Gln Arg Tyr Ile Tyr Glu Val Glu Gln
218      245      250      255
220 Gly Leu Tyr Pro Ala Glu Glu His Ser Phe Gln
221      260      265
224 <210> SEQ ID NO: 9
225 <211> LENGTH: 349
226 <212> TYPE: PRT
227 <213> ORGANISM: Aspergillus nidulans
229 <400> SEQUENCE: 9
230 Met Thr Phe Leu Arg Ile Ala Thr Lys Arg Ala Ile Tyr Leu His Arg
231 1 5 10 15
233 Pro Ala Asn Pro Ala Leu Pro Thr Ser Ser Ile Leu Pro Val Leu His
234 20 25 30
236 Ser Thr Asn Val Ala Thr Arg Val Pro Ser Pro Cys Ala Ile Arg His
237 35 40 45
239 Ser Ser His Ser Pro Leu Gly Ala Ala Gln Ala Asn Pro Arg Lys Lys
240 50 55 60
242 Val Thr Met Gln Thr Leu Arg Asn Leu Tyr Lys Lys Gly Glu Pro Ile
243 65 70 75 80
245 Thr Met Leu Thr Ala His Asp Phe Pro Ser Ala His Val Ala Asp Ala
246 85 90 95
248 Ala Gly Met Asp Met Ile Leu Val Gly Asp Ser Leu Ala Met Val Ala
249 100 105 110
251 Leu Gly Met Gln Asp Thr Ser Glu Val Thr Leu Asp Asp Met Leu Val
252 115 120 125
254 His Cys Arg Ser Val Ala Arg Ala Ala Gln Ser Ala Phe Thr Val Ser
255 130 135 140
257 Asp Leu Pro Met Gly Ser Tyr Glu Val Ser Pro Glu Gln Ala Leu Gln
258 145 150 155 160
260 Ser Ala Ile Arg Ile Val Lys Glu Gly Arg Val Gln Gly Val Lys Leu
261 165 170 175
263 Glu Gly Gly Glu Glu Met Ala Pro Ala Ile Lys Arg Ile Thr Thr Ala
264 180 185 190
266 Gly Ile Pro Val Val Gly His Ile Gly Leu Thr Pro Gln Arg Gln Asn
267 195 200 205
269 Ala Leu Gly Gly Phe Arg Val Gln Gly Lys Ser Thr Thr Asp Ala Leu
270 210 215 220
272 Lys Leu Leu Lys Asp Ala Leu Ala Val Gln Glu Ala Gly Ala Phe Met
273 225 230 235 240
275 Ile Val Ile Glu Ala Val Pro Pro Glu Ile Ala Ser Ile Val Thr Gln
276 245 250 255
278 Lys Leu Ser Val Pro Thr Ile Gly Ile Gly Ala Gly Asn Gly Cys Ser
279 260 265 270
281 Gly Gln Val Leu Val Gln Ile Asp Met Thr Gly Asn Phe Pro Pro Gly
282 275 280 285
284 Arg Phe Leu Pro Lys Phe Val Lys Gln Tyr Ala Asn Val Trp Asn Glu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/820,745

DATE: 11/26/2001

TIME: 12:34:34

Input Set : A:\620-139.app

Output Set: N:\CRF3\11212001\I820745.raw

```

285      290      295      300
287 Ala Leu Gln Gly Ile Gln Gln Tyr Arg Glu Glu Val Lys Ser Arg Ala
288 305      310      315      320
290 Tyr Pro Ala Glu Gln His Thr Tyr Pro Ile Pro Lys Glu Glu Leu Val
291      325      330      335
293 Glu Phe Gln Lys Ala Val Asp Glu Leu Pro Glu Glu Lys
294      340      345
297 <210> SEQ ID NO: 10
298 <211> LENGTH: 347
299 <212> TYPE: PRT
300 <213> ORGANISM: Arabidopsis thaliana
302 <400> SEQUENCE: 10
303 Met Ala Ser Ser Leu Thr Arg Asn Cys Ser Arg Phe Ser Lys Ala Ile
304 1 5 10 15
306 Ser Val Arg Phe Met Ser Asn Leu Pro Glu Asn Thr Val Tyr Gly Gly
307 20 25 30
309 Pro Lys Pro Gln Asn Pro Asn Gln Arg Val Thr Leu Thr His Leu Arg
310 35 40 45
312 Gln Lys His Arg Arg Gly Glu Pro Ile Thr Val Val Thr Ala Tyr Asp
313 50 55 60
315 Tyr Pro Ser Ala Val His Leu Asp Thr Ala Gly Ile Asp Val Cys Leu
316 65 70 75 80
318 Val Gly Asp Ser Ala Ser Met Val Val His Gly His Asp Thr Thr Leu
319 85 90 95
321 Pro Ile Ser Leu Asp Glu Met Leu Val His Cys Arg Ala Val Ala Arg
322 100 105 110
324 Gly Ala Lys Arg Pro Leu Leu Val Gly Asp Leu Pro Phe Gly Thr Tyr
325 115 120 125
327 Glu Ser Ser Ser Ser Gln Ala Val Asp Thr Ala Val Arg Val Leu Lys
328 130 135 140
330 Glu Gly Gly Met Asp Ala Ile Lys Leu Glu Gly Gly Ser Ala Ser Arg
331 145 150 155 160
333 Ile Thr Ala Ala Lys Ala Ile Val Glu Ala Gly Ile Ala Val Ile Gly
334 165 170 175
336 His Val Gly Leu Thr Pro Gln Ala Ile Ser Val Leu Gly Gly Phe Arg
337 180 185 190
339 Pro Gln Gly Arg Asn Ile Ala Ser Ala Val Lys Val Val Glu Thr Ala
340 195 200 205
342 Met Ala Leu Gln Glu Ala Gly Cys Phe Ser Val Val Leu Glu Cys Val
343 210 215 220
345 Pro Pro Pro Val Ala Ala Ala Ala Thr Ser Ala Leu Lys Ile Pro Thr
346 225 230 235 240
348 Ile Gly Ile Gly Ala Gly Pro Phe Cys Ser Gly Gln Val Leu Val Tyr
349 245 250 255
351 His Asp Leu Leu Gly Met Met Gln His Pro His His Ala Lys Val Thr
352 260 265 270
354 Pro Lys Phe Cys Lys Gln Tyr Ala Asn Val Gly Glu Val Ile Asn Lys
355 275 280 285
357 Ala Leu Met Glu Tyr Lys Glu Glu Val Ser Lys Lys Val Phe Pro Gly

```

Use of n and / or Xaa has been detected in the
 Sequence Listing. Review the Sequence Listing
 to ensure a corresponding explanation is present
 in the <220> to <223> fields of each sequence
 using n or Xaa.

VERIFICATION SUMMARY

DATE: 11/26/2001

PATENT APPLICATION: US/09/820,745

TIME: 12:34:35

Input Set : A:\620-139.app

Output Set: N:\CRF3\11212001\I820745.raw

L:62 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:486 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:489 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:492 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:495 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:498 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:501 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:504 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:507 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:510 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:513 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:516 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:519 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:522 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:525 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:528 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:531 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:534 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:537 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12